SUPPLEMENTARY INFORMATION

Healthy participant (phase I model) popPK dataset creation

The integrated analysis datasets included treatment allocation, actual dosing dates and times (and actual dose), actual and nominal pharmacokinetic (PK) sampling dates and times, derived actual and nominal time after dose (and time after first dose), letermovir plasma concentrations, and participant demographic data.

The anonymity of participants who provided genetic data was protected by double coding, such that there was no allocation number (AN), and only genetic identification numbers (ID) were available. An AN-genetic mapping key file was used to connect the genetic data with PK data and was only accessible by designated personnel. The modeler did not have access to the AN-genetic ID (and thus genetic data) mapping. This dataset included genetic variations in *OATP1B1* (variants rs4149056 and rs2306283) and *UGT1A1* (variant rs4148323) with dummy participant identifiers.

No transformations of the source data were performed. Missing data and outliers were excluded from the population analysis. All PK observations without recorded sampling times and dates, or which were not associated with a dosing event, were regarded as non-evaluable. Observations below the lower limit of quantification (LLOQ; <1 ng/ml) were retained in source data, but excluded from the analysis. Flags were included in the dataset to identify any excluded data points or concentration below the LLOQ to allow for obtaining post hoc predictions at these time points. The effects for the following covariates were evaluated on clearance (CL) and volume of distribution (sum of V_1 [central volume of distribution], V_2 [first peripheral volume of distribution], V_3 [second peripheral volume of distribution], and V_4 [third peripheral volume of distribution]) (V_d): age; body weight; gender; race; ethnicity; *OATP1B1* genotype; and *UGT1A1* genotype. Covariates were excluded from the population analysis if >30% of participants did not provide a value. Otherwise, the value of missing covariates was imputed as the median of the remaining values from an appropriate sub-population. For categorical covariates, such as race and gender, the most frequent occurring value was imputed for each trial. Data with population conditional weighted residual (CWRES) >4 or individual weighted residual >4 were considered potential outliers.

Healthy participant (phase I model) exploratory data analysis

Prior to the population PK (popPK) analysis, observed letermovir concentrations were plotted against time and stratified by key trial design elements (e.g., treatment or regimen). These exploratory plots were used to inform selection of a starting structural model and inform covariate selection.

Deriving healthy participant (phase I model) letermovir-simulated exposure

Simulations were performed with the model in 1000 typical participants per dose level to visualize the expected-dose nonlinearities over the full dose range of 30–960-mg single- or multiple-dose letermovir, after oral or i.v. administration, and to evaluate the impact of identified covariate effects on letermovir exposure after multiple oral dosing of 480 mg. The PK parameters estimated were the area under the concentration-time curve from time 0 to infinity ($AUC_{0-\infty}$), area under the concentration-time curve between 0 and 24 h postdose at steady-state (AUC_{0-24}), and maximum letermovir concentration (C_{max}).

Healthy participant (phase I model) covariate analysis

For the healthy participant (phase I) model, the following covariates were tested: age, weight, gender, race, ethnicity, *OATP1B1* genotype, and *UGT1A1* genotype.

HSCT recipient (phase III model) popPK source data

The phase III trial was a randomized, double-blind, placebo-controlled study that assessed the efficacy and safety of 480 mg/day oral and i.v. letermovir (or 240 mg/day with concomitant cyclosporine A) in cytomegalovirus (CMV)-seropositive hematopoietic stem cell transplant (HSCT) recipients (clinicaltrials.gov, NCT02137772). The primary endpoint was clinically significant CMV infection through week 24 post-HSCT.[11] Letermovir efficacy and safety were also assessed in CMV-seropositive HSCT recipients in the phase II trial (clinicaltrials.gov, NCT01063829).[8] The phase I trials assessed the safety and PK of letermovir in healthy participants, including one study conducted in an Asian population, and were included to enrich the sparse phase IIb and III data. Since the objective was to describe steady-state PK, only data obtained following a minimum of 1 week of dosing and less than 72 h after the last dose were included from all trials. Steady-state phase I data were included only from studies with the same dosing schedule as used in the phase III trial. In the

phase I trial, full PK profiles were obtained following 1 week of letermovir treatment; in the phase IIb trial, sparse PK data were collected weekly. Rich phase III data were obtained from 74 participants following 1 week of treatment where multiple PK samples were obtained following one dosing occasion; sparse samples were obtained from the remaining participants pre-dose at weeks 2, 4, 6, 8, 10, 12, and 14.

HSCT recipient (phase III model) covariate analysis

Covariates of clinical interest were pre-specified for analysis. For CL and bioavailability, covariate relationships were tested for age, body weight, creatinine clearance (CrCl), aspartate transaminase (AST), alanine transaminase (ALT), alkaline phosphatase (ALP), albumin, tacrolimus, sirolimus, posaconazole, voriconazole, ciprofloxacin, fluconazole, amlodipine, prednisone, budesonide, midazolam, fentanyl, azithromycin, sex, race, and Asian origin. Covariate relationships for age, body weight, CrCl, AST, ALT, ALP, albumin, sex, race, and Asian origin were pre-specified for testing on V_1 and V_2; however, as interindividual variability (IIV) was not quantified for V_1, these covariate relationships were tested only for V_2.

Model diagnostics and qualification (healthy participants [phase I model] and HSCT recipients [phase III model])

Diagnostic plots of observed data vs. population predictions (PRED) and individual prediction were examined for adequate fit. Plots of CWRES vs. PRED and vs. time (time after last dose as well as time after first dose) were inspected for evidence of systematic lack of fit, and to confirm the absence of bias in the error distributions.

Individual deviations (ETAs) from the population mean are expected to be normally distributed with mean zero. To verify absence of bias, IIVs were graphed in scatter plots vs. key continuous model covariates with potential trends visualized with a Loess smoother, and in a box and whisker plot vs. key categorical model covariates. Visual predictive checks (VPCs) were constructed stratified on dose and route of administration in order to evaluate the model's prediction performance.

For the healthy participant popPK model, parameter precision for the popPK model was evaluated using the standard error calculated from nonlinear mixed effects modeling (NONMEM) and non-parametric bootstrap analysis. A bootstrap procedure was carried out to obtain non-parametric

standard errors for parameter estimates. Using the bootstrap procedure as implemented in Perl-speaks-NONMEM (PsN), the model dataset was resampled (with replacement) 500 times to derive 90% confidence intervals (CIs) for parameter values. The resampling was stratified by trial, such that each trial was, on average, equally represented in the resampled datasets. The model was then fitted to these resampled datasets providing 500 sets of parameter estimates that collectively formed the posterior parameter estimate distribution. The bootstrap was carried out for the final model only, and the bootstrapped 90% CIs of the parameter estimates were reported.

HSCT recipient (phase III) model: deriving letermovir individual exposure

Exposure in terms of area under the concentration-time curve from 0 to 24 hours postdose at steady-state (AUC_{ss}), C_{max}, and 100-fold difference in minimum concentration (C_{trough}) at steady-state was calculated for each subject using individual predicted post hoc PK parameters obtained from the HSCT recipient (phase III) population PK model and the individual dosing history without interoccasion variability (IOV). The predicted median AUC_{ss} was generated for four dosing regimens: oral and i.v. administration of letermovir, with and without co-administration of CSA. Phase III trial participants received a complex dose regimen that, for some individuals, included all four of these dosing regimens. Individual exposure predictions for all phase III participants were calculated as the weighted average exposure based on the dosing history for each individual.

Deriving letermovir-simulated exposure

For the healthy participant (phase I) model, simulations were performed in 1000 typical participants per dose level to visualize the expected-dose nonlinearities over the full dose range of 30–960 mg single- or multiple-dose letermovir, after oral or i.v. administration, and to evaluate the impact of identified covariate effects on letermovir exposure after multiple oral dosing of 480 mg. The PK parameters estimated were $AUC_{0-\infty}$, AUC_{0-24}, and C_{max}.

Using the HSCT recipient (phase III) model, letermovir exposure in the general HSCT population was estimated following the four different dosing regimens by performing simulations with 1000 subjects per dosing regimen without residual variability and IOV. Aggregate statistics (median and prediction interval) were then used to characterize typical concentration-time profiles and associated PK parameters of AUC_{ss}, C_{max}, and C_{trough}.